

Selection of control diagnostic symptoms and technical state features of marine engine conventional injectors using multiple regression analysis

Jan Monieta^{1,*} , Lech Władysław Kasyk¹ 

¹ Maritime University of Szczecin, Poland

* Corresponding Author: jan.monieta@gmail.com

Received: 15 January 2026

Revised: 28 September 2026

Accepted: 9 October 2026

Online: 23 October 2026

This is an open access article
under the [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Abstract

This article presents the results of a study of multivariate application statistics to approximate experimental test results. The influence of explanatory variables on the approximation using multiple regression was assessed using various quantities that affected the test conditions and technical state. Measurements of key injector performance features were made using various geometric and fuel dose measurement instruments. The diagnostic symptoms of the injector nozzle vibration displacement signals in the frequency domain, which are most correlated with relevant properties of the technical state and test conditions, were selected. The relevance of individual state features was tested by best fitting the diagnostic parameter model to the dependencies of the variables, including features, and the control quantities. The most useful symptom was the first component of the envelope amplitudes in the nozzle needle displacement spectra, and the most relevant properties of the injection apparatus's technical state were the injected fuel dose at the exact injection frequency. In this way, a global technical state feature was selected for quality control during the production and operation phases.

Keywords: marine injectors, diagnostic symptoms, multiple linear regression, symptoms selection, controlled state features

Article citation:

Monieta J, Kasyk LW, Selection of control diagnostic symptoms and technical state features of marine engine conventional injectors using multiple regression analysis, *Eksploracja i Niezawodność – Maintenance and Reliability* 2027; 29(2) [http://doi.org/10.17531/ein/241609](https://doi.org/10.17531/ein/241609)

Highlights

- Most correlated diagnostic of needle vibration displacement signals in the frequency domain.
- Diagnostic parameters were related to relevant features of the injector's technical state.
- Multiple regression was used to select the individual state features.
- This has applications for the choice of state features after production or maintenance.
- Useful symptoms are the first component of the envelope amplitudes of the nozzle displacement.

1. Introduction

Data analysis is used to determine the final experimental results, their measurement errors, uncertainties, and metrology limitations, which determine the planning and implementation of the research. At the same time, it refers to mathematical modeling by developing a qualitative model and then object

functions that can be transformed into a mathematical model of the internal combustion engine [1].

Forecasting in marine engine operations has been applied to:

- approximation of diagnostic symptom values,
- predicting the technical state in a specific future,
- determining the time to failure.

In many studies, a decomposition level or a polynomial level is assumed based on literature or intuition. Therefore, an attempt was made to study the effect of the level of approximation on the diagnostic measure values [2]. Methods were used to develop results from diagnostic experiments, employing statistical techniques such as correlation, regression, factor analysis, and stochastic approximations [3,4].

A mathematical model is expected to be fully logically adequate for the object, and an object function is expected to be sufficient in the study area for the test results. The best function is consistent with the model and approximates the results of the

measurements in the study area of the input quantities well. To implement the task, it is necessary to select the function of the research object and determine how to approximate the measurement results using that function.

Some internal combustion engine tests are carried out in accordance with experimental theory [5]. Modern versions of computer programs for analyzing experimental data sometimes use multiple regression models, which facilitate the study of interdependencies and correlations between model input parameters and a single output variable. It is also used to prepare experimental material for other advanced diagnostic tools, such as neural network models, which, when properly configured, enable the analysis of measurement results recorded during dynamic processes.

This article presents the results of using multiple regression to assess the influence of state properties and control conditions on selected diagnostic parameters of vibration signals in the most helpful frequency domain, using the theory of experimental design and static plan-determined selection multiple. Optimal signal analysis methods and diagnostic symptoms correlated with significant properties of the technical state were sought.

2. Background

2.1. State of knowledge

The construction of effective diagnostic models for multidimensional, complex objects [6,7], such as reciprocating internal combustion engines, has been subject to considerable difficulties when using active feedback-based approaches [8–10]. The main reasons for this are the varying operating conditions of internal combustion engines, as well as masking of diagnostic information, the difficulties in achieving certain states of wear or degree of component damage, the difficulties in the propagation paths of diagnostic signals from damage locations to reception points in a complex internal combustion engine and test bench, etc. [11]. In regression analysis, statistical tests aim to explain the relationship between different population characteristics under study [12]. A population is understood to be a set of elements having a constant characteristic that unites them and distinguishes them from other sets.

Approximation involves assessing imprecision, statistical

correlations, and evaluating the adequacy of an object function in relation to measurement results or a mathematical model. It is a method of substituting, with the required accuracy a set of values of the output quantities $\{y\}$, being the results of measurements, corresponding to a set of input quantities $\{x\}$, defined by a relation:

$$\tilde{y} = \tilde{f}(x) \sim y = f(x) \quad (1)$$

The most important approximation methods include [13] least squares, interpolation, alignment, average, and frequency analysis. When applied to probabilistic models, approximation methods such as those for deterministic models can be used by treating the output averages ($y \equiv y_a$) as determined quantities, or special methods of random process analysis can be applied.

In the context of technical equipment diagnostics, multiple regression can be used to understand factors affecting the performance or state of an object [14,15]. By identifying relevant independent variables (such as device parameters, environmental conditions, or usage patterns) and their respective contributions to the dependent variable (e.g., device performance, damage rate, or diagnostics), scientists and engineers can develop more accurate and reliable diagnostic models.

The article by Lin et al. [16] converts a simple fuzzy linear regression model into a multiple fuzzy regression model. Least-squares methods were used to estimate parameters using D_K metrics and stepwise regression. Bootstrap methods were used to estimate standard errors and determine statistical test values. Fuzzy diagnostics using D_K -metrics were also used to detect outliers.

Few works include multiple linear regression profile analyses [17–20]. Ćwikła and Paprocka [17] used a multiple linear regression model to describe the impact of historical data on pump performance for maintenance purposes. By identifying the key factors affecting pump performance, the authors optimized maintenance activities and improved equipment reliability.

Mahmoud [21] presented an alternative approach to the analysis using linear phase regression. In the solution, the model response was monitored using intersection, slope, and variance. Using simulations, the performance of the developed method was compared with the performance of existing methods for monitoring data from multiple linear models in terms of signal

waveform probabilities. The advantage of this method over existing methods was the significantly improved detection of changes in the parameters of linear models within a multivariate region.

Galindo et al. [22] presented a detailed methodology for determining heat release in two-stroke engines over a wide range of operating conditions: the empirical data obtained were used to construct a Wiebe function, in which four parameters were correlated with engine parameters obtained from a 1D simulation. Multiple regression analysis showed that charge density, residual gas fraction, spark ignition time, and average piston velocity were significant engine parameters that affect Wiebe parameters.

According to Sheikhi et al. [23], one important assumption in multiple regression is the independence of the explanatory variables. The regression equation was used to diagnose heteroskedasticity, and the chosen regression model was also tested using simulated data. A cross-validation method was performed to achieve unbiased results.

Adamkiewicz [24] applied multiple regression to predictive estimation, forecasting the technical state of marine internal combustion turbine engines. Multiple regression describes more complex models, e.g., polynomial models of different degrees and non-linear models that are reducible to linear models (product models, exponential models, fractional models, etc.) [11]. It was possible to transform the values of variables in a multi-element sample with random loadings into conditions comparable to the loadings of the parameter diagnostic signal, using the model parameters. The notes that the advantages of regression models, as opposed to polynomial, power, and other models, include interpreting their coefficient and the values of their high correlation with experimental results.

Rudnicki and Zadrąg [25] used multivariate models to facilitate the analysis of measurement results. These models allowed simultaneous assessment of effects and relationships between multiple output variables. The authors demonstrated the advantages of using analytical tools and an exemplary application of a neural model developed on a marine internal combustion engine bench.

Fault diagnosis techniques for marine compression-ignition engines are essential in studying the safety and reliability of marine vessel operations [26,27]. The literature indicates that

single faults is easy, whereas diagnosing multiple faults simultaneously is rarely successful. In this study, diagnostic methods based on parametric thermal analysis combined with various neural network algorithms were developed and applied to diagnose multiple faults in a marine self-ignition engine. The authors showed that the Levenberg-Marquardt backpropagation neural network achieved the highest diagnostic accuracy, reaching 100% for single failures and 88.89% for multiple failures. The Bayesian backpropagation regularization neural network achieved a diagnostic accuracy rate of 100% for single faults and 55.56% for multiple faults as reported in [27].

Several studies have demonstrated the effectiveness of multiple regression techniques in diagnosing various technical states in complex systems. For example, Ploeg et al. [28] discussed the use of regression-based algorithms to estimate multiple faults in dynamic systems, highlighting their tractability and performance limits. Similarly, Lim et al. [29] proposed a method for the rapid diagnosis of multiple faults using an efficient candidate selection algorithm that outperformed other established methods for damage detection and isolation.

Recent analyses have shown that multiple forms of unfitness are often present in disability systems. Therefore, the diagnosis of multiple damages is vital. Lim et al. [29] propose a method for multiple-state diagnostics, using a selection algorithm that can capture different damage behaviors.

Antory et al. [30] used a non-linear variable reconstruction algorithm and two-dimensional scatter plots to isolate damages. The non-linear variable description can remove the condition signature from the selected measurement data, facilitating the identification and isolation of the cause of the initiation of an internal combustion engine response. Artificial neural networks and support vector machines, due to their accuracy and suitability for analyzing non-linear issues, have been applied in many research disciplines [31].

The problem of detecting damage to marine internal combustion engines [32], where diagnostic symptoms exceed limit values, has been partially solved. The difficulties associated with the occurrence of false alarms and second-type errors in recognizing a fault state, are an ability that arise when continuously monitoring the self-ignition condition of an internal combustion engine.

Wang et al. [33] proposed a multivariate statistical fault detection approach to monitor engine faults with weak signatures. In this method, they first used principal component analysis to project the engine observations into a principal component subspace and a residual subspace.

In diagnostics, during fault location, the current parameter values of the diagnostic signals are compared with the values of the reference signatures. After all symptoms of the i -th fault have occurred, the diagnostic system can conclude that the fault has occurred [18,34]. Damages are difficult to distinguish when their signatures are the same. However, selective signatures, using subsets of signal measures, can be used to locate damage without losing damage identification.

Damage to an object can be described by signatures $v_i(\tau, \theta)$, which are a dynamic transformation by the working motion in the machine and the dynamic time τ , at the point of origin of the unfitness state with vector coordinate r_i . It is possible to perform dynamic transformations of unfitness in signal [35]:

$$F_i(\theta) \Rightarrow v_i(\tau, \theta, r_i) \quad (2)$$

where θ is the operating time of the object.

Hotelling's T^2 and Q statistics were introduced to detect deviations of engine observations into the principal component subspace [36]. The T^2 and Q statistics were developed by considering the correlations among various parameters [37] to detect errors with weak signatures. The proposed approach is demonstrated using a marine Diesel engine. The results showed that this approach can effectively detect internal combustion engine damage with weak signatures.

Jafari et al. [38] combined a correlation matrix and hierarchical clustering to assess the strength of relationships between diagnostic parameters from different groups of marine self-ignition engines. Some monitored parameters, such as lubricating oil temperature, should, by design, be constant. The authors proposed another method that uses probability density functions for the diagnostic parameters with the weakest relationships and the greatest variability.

The Sehili [39] study included numerical modeling of injector temperatures to enable spatial and temporal monitoring under various operating conditions. Temperature changes in conventional dual-fuel injectors were compared. However, limitations in installing temperature sensors were noted.

Another application of multiple regression in technical

equipment diagnostics was in the state monitoring of aircraft gas turbine engines. Abdullayev et al. [40] discussed the use of fuzzy multiple linear and non-linear regression models to diagnose the state of aircraft gas turbine engines, particularly in the early stages of damage development, when available information is limited and uncertain.

Despite advances in marine engine diagnostic technologies and systems, detecting damage combinations over the entire range of operation is very difficult [41]. The authors' aim was first to investigate the method of logarithmic path signatures for creating information and dimensionality reduction from pressure waveform signals in individual cylinders and to select the most effective hybrid method. The results showed that the use of principal logarithmic signatures and a regression model, with a two hidden layer architecture with ten neurons in each, was the most promising in terms of mean squared error.

In his publications, Bejger [42] presented studies using acoustic emission to diagnose the injection subsystem of marine engines, including injectors. However, he did not use diagnostic parameters or link acoustic emission analyses to graded changes in technical state.

The study included numerical modeling of injector temperatures to enable spatial and temporal monitoring under various operating conditions [39]. Temperature changes in conventional dual-fuel injectors were compared.

Research on the diagnostics of marine diesel engine injectors is currently focused on non-invasive, real-time monitoring and the transmission of data to shore using artificial intelligence [2,43]. The development of non-invasive sensors is limited, and the analysis is not fully optimized to identify specific faults without disassembling the component. Existing artificial intelligence models require refinement to handle nonlinear and nonstationary signals generated by the modern electronically controlled engines [43]. Despite progress, significant research gaps remain due to the extreme operating conditions of injectors, complex fuel types, and dual-fuel designs [44], as well as difficulties in accessing injectors in running engines.

Injectors in modern common-rail systems of compression-ignition engines were also diagnosed. Ke et al. [45] developed an intelligent diagnostic method using Composite Hierarchical Dispersion Entropy with Grasshopper optimization. The authors

proposed the IGOA-LSSVM multi-class classifier for pattern recognition, achieving higher recognition accuracy than competing methods.

2.2. Approximation in the planning of experiments

The choice of the approximating function is not straightforward,

$$y = b_0 + b_{11}x_1 + b_{21}x_2 + b_{31}x_3 + \dots + b_{i1}x_i + b_{12}x_1^2 + b_{22}x_2^2 + \dots + b_{i2}x_i^2 + b_{1k}x_1^k + b_{2k}x_2^k + \dots + b_{ik}x_i^k \quad (3)$$

In research, a compromise is made by allowing nonlinearity in the actual test objects by increasing the polynomial degree. For plan efficiency, the polynomial degree is decreased [46]. However, each additional coefficient that increases the number of unknowns in the polynomial coefficients adds an extra set of input values to the test plan, and each new set may be a new version of the prototype object, which is a complicating factor. Point estimates of the regression coefficients, including their direction and statistical significance, were determined in the regression model [12].

The minimum number of outputs for systems of input values is 3 for polynomial approximation [13]. The main problem in developing the results of passive experiments is determining the form of the equations and their coefficients. The equations with the highest coefficient of determination [5] between the variables were adopted, and when the correlation coefficients were close, a lower-degree model was chosen.

Modifications restrict the degree of the polynomial to be linear, linear in double products, second-degree, quadratic with first-order coefficients, etc. An algebraic or trigonometric polynomial is used to approximate measurement results as an object function. Thus, there are grounds for considering the polynomial form as one possible form of the mathematical model of an internal combustion engine.

Xu et al. [47,48] stated that the expert system outperforms artificial neural network models, support vector machine models, with binary logistic regression model and binary logistic regression models in terms of accuracy and stability, and can identify concurrent damages.

Li and Wang [7] used the study of four models: multiple linear regression, stepwise regression, backpropagation neural network, and extreme learning machine model to simulate dam deformation. The results showed that the multiple-point model had higher prediction accuracy than the single-point model.

as its form does not always follow from the experimentally determined relationship. Most experimental plans, including selection plans, are associated with an approximation by a second or higher degree. In the general case of a multi-input object, such as a marine internal combustion engine, the most common form of the polynomial is:

In optimization studies, quadratic programming is non-linear, and a simpler quadratic polynomial can approximate polynomials of higher degrees than two, thereby avoiding a complex mathematical task. In studies with multiple input quantities, a second-degree polynomial is often adopted, which can be reduced to a linear form via linearizing transformations.

Previous studies have dealt sparsely with the application of vibration signals, including vibration nozzle needle displacements [49,50], to multi-correlated diagnostics. Their work has focused on other areas of signal analysis, and multi-correlated regression has not been used in such a comprehensive statistical approach.

Most diagnostics are currently performed during maintenance or when symptoms occur, rather than through continuous, automated on-board monitoring. Signal acquisition difficulties in self-ignition engines are often compounded by injectors located under valve covers, which limit sensor placement. Many diagnostic models are validated under laboratory conditions. Still, their accuracy is limited in real-world conditions, including likelihood diagnostics that can distinguish between injector degradation and temporary performance issues caused by variable fuel quality or operating conditions. Given the complexity of a dual-fuel system with engines switching between low- and high-pressure gas injection, diagnostics must be adapted to monitor both the diesel-engine pilot injectors and the gas-supply systems [39].

The optimal innovative location of the vibration sensor and the direction of sensor mounting on the injection line, as well as the selection of useful diagnostic symptoms, are presented in [51]. It has been demonstrated that vibration acceleration signals recorded along the injection pipe have the highest energy. By utilizing multiple parameters for a given signal analysis field and selecting cumulative diagnostic symptoms. As a result, the tested fuel injection subsystem is classified as

pass or fail, using unprecedented tolerance ranges based on the mean value for the pass condition: ± 3 times the standard error of that mean.

The literature was dominated by partial works that did not link diagnostic symptoms to measurements of graded condition features. The authors of this work aimed to fill a gap in existing state knowledge.

3. Methods and measures

3.1. Objects of study

Injectors are the most unreliable assemblies in marine reciprocating internal combustion engines in studies by the author [5] and other authors [19], so they are the objects of study. The engines studied were primarily four-stroke medium-speed turbocharged engines. The technical data of the type of internal combustion engine tested are shown in Table 1, from which a wide range of tests can be obtained. The most frequently tested engines were the 6AL20/24, designed by the same licensor and manufactured by the same manufacturer. These were engines with direct fuel injection into the cylinders, driven by a camshaft in the engine block, supercharged by a turbocharger, with classic injection subsystems.

The technical data of the most tested marine engine is shown in Table 1.

Table 1. Technical data of the most frequently tested marine internal combustion engines.

Engine type	6AL20/24D(H)
Construction type	in-line
Number of cylinders	6
Direction of rotation	right-hand
Ignition order	1 – 4 – 2 – 6 – 3 – 5
Rated camshaft rotational speed [rpm]	750
Average piston velocity [m/s]	6
Cylinder diameter [mm]	200
Piston stroke [mm]	240
Stroke volume [L]	7.540
Cylinder rated power [kW]	72.5
Compression ratio	1:12.7
Mean effective pressure [MPa]	1.48
Injector opening pressure [MPa]	24.5 or 40
Max. combustion pressure [MPa]	10.0; 12.5

D – fuelled with distillate fuel, H – fuelled with residual fuel.

The reliability tests showed that most of the damage occurred during the initial period of injector operation, indicating initial unsuitability [5]. Insufficient quality control was identified during manufacturing and maintenance stages, necessitating diagnostic checks. Verification tests were carried

out at the Laboratory of Marine Power Plants of the Maritime University of Szczecin and on sea-going vessels.

3.2. Research method

The research was conducted using measured data collected in experiments, simultaneously recording the operating conditions of the self-ignition engine or injection subsystem. In experimental studies, it was necessary to conduct measurements under reproducible conditions, to search for optimal sensor mounting locations, to select analysis bands, and minimize signal interference [2,51].

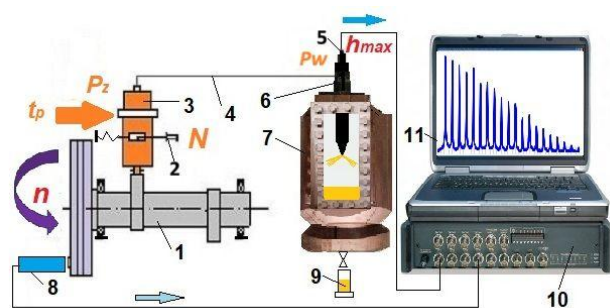
The tests were first conducted on a test bench with a separate injection subsystem outside the engine, consisting of actual AL20/24 engine components. Fig. 1a illustrates the fuel injection and metering stand, as well as the signal acquisition and processing (Fig. 1b). Fuel injection took place into a chamber at ambient temperature with measurement of the injected fuel doses (Fig. 1b) [51]. The test conditions and input values were determined based on operational tests [1].

Among others, a computer program, DaqView from IOtech [2], was used for signal acquisition, and the Signal Analysis System [51] was used for signal analysis. DaqView is a configuration and measurement software that works with measurement systems from various companies. It allows devices to be configured and measurements to be taken. The software will enable users to select settings and view real-time measurement results. The measured signals were pre-analyzed using the software for measurement control, signal processing, and export [2,52,53].

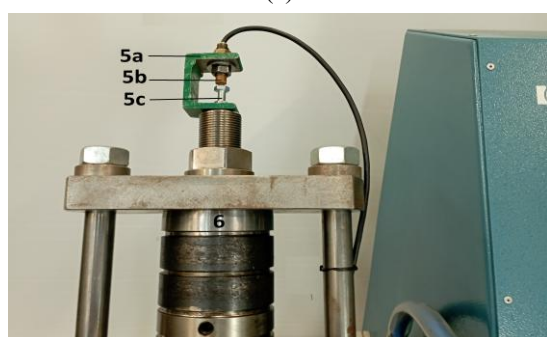
The signal processing channels consisted of CFR-20 photo-optical camshaft position sensors and DT10 eddy current vibration displacement sensors, along with a set of amplifiers and matching circuits. These components amplified the signal according to the sensors' sensitivity and adapted it to the inputs of the analog-to-digital converters' inputs. Before starting the measurements, the environmental parameters were brought to reference conditions, i.e., the room temperature was set to 18–22°C, and the relative humidity to 50–70%. The photoelectric sensor signals served as a time marker. Another element of the analog path was a low-pass filter, which also served as an anti-aliasing filter [53].

The measurement signals, after passing through the anti-

aliasing low-pass filter [53], were fed into the analog-to-digital converter (ADC) inputs. The signals were analyzed using a desktop or laptop computer. Preliminary signal processing was performed using time selection, filtering, synchronous averaging, and envelope detection, among other methods.



(a)



(b)

Figure 1. Diagram of the injection subsystem with a measuring system (a) and a view of the signal acquisition and processing system (b): 1 – camshaft, 2 – fuel regulating linkage rod, 3 – injection pump, 4 – injection high pressure pipe, 5 – eddy current displacement vibration sensor, 5a – holder, 5b – displacement sensor, 5c – lifting mandrel, 6 – injection valve, 7 – fuel injection chamber, 8 – camshaft position sensor, 9 – fuel dose measuring cylinder, 10 – connection terminal, 11 – laptop.

The signal analysis system (SAS) is a program that controls the ADC card, and enables fundamental analyses in the domain. The system enables the selection of gain (including smooth gain), frequency range, number of samples for further analysis, number of averages, one of the signal channels, trigger type (continuous or with voltage acquisition trigger in the trigger channel), trigger channel, trigger level, time delay, and averaging mode. Some of the analysis results were transferred to Excel, Statistica, and MATLAB for further processing, graphical presentation, and selecting the most useful symptoms, etc.

The technical state of the tested injectors was determined by geometric measurements of relevant state features [54]. Fig. 2a shows a view of the microscope with a computer and software, and Fig. 2 shows a station for measuring internal dimensions with a self-centering digital sensor diameter gauge with an expansion tip. The diameter gauge was mounted on a measuring stand with adjustable lifting height. Diameter measurements were taken in three planes perpendicular to the longitudinal axis and in four cross-sections spaces every 45° through their axes.



(a)



(b)

Figure 2. View of the microscope with a computer and software for geometric measurements (a) and a station for measuring the diameter of the guide part of the injector nozzle body (b): 1 – stand, 2 – nozzle body, 3 – self-centering diameter measuring bore gauges, 4 – floating holder, 5 – digital inductive sensor.

The architecture of a computer-based measurement system that provides sensor support for power supply, signal loading, synchronous or asynchronous analogue-to-digital processing, preprocessing of measured analogue signals, analysis of selected signals and their archiving in arrays, and development of approximation models is shown in Fig. 3.

The analysis of the signals of the working and associated process parameters was carried out in the time domain, amplitude, frequency, and time-frequency domains; it also

decimated, windowed, wavelet statistics, color analysis, and state classification was carried out [2].

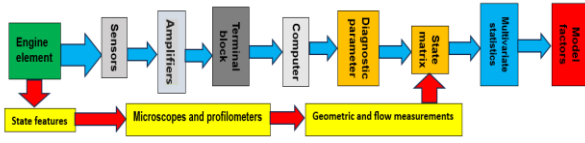


Figure 3. Methodology architecture from study object to regression model development.

During the fuel injection process, it is possible to determine the frequencies of phenomena related to fuel flow and the behavior of injector components. Sudden changes in fuel pressure cause hydraulic shocks in the fuel pipe as the injector needle is raised and lowered during the pumping process. This phenomenon causes vibrations in the fuel pipe with a frequency of:

$$f_{pp} = in \quad [\text{Hz}] \quad i = 1, 2, \dots \quad (4)$$

where: i – successive harmonic, n – camshaft rotational speed [1/s].

Characteristic frequencies in the vibrations during fuel injection were calculated to identify the components in the spectra of pressure in the injection pipe and the displacements and accelerations of the injector needle. The needle dynamics equation describes the movement of the injector needle during fuel injection:

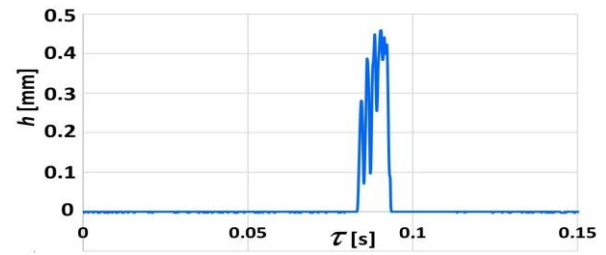
$$m_w \frac{d^2 h}{d\tau^2} + A_g \frac{dh}{d\tau} + k_{sw}(h + h_{wo}) - p_i(A_{ip} - A_i) - p_s A_i = 0 \quad (5)$$

where: A_g – flow area through the needle seat, A_i – cross-sectional area of the needle, A_{ip} – cross-sectional area of the needle guide, h – instantaneous displacement of the nozzle needle, h_{wo} – initial tension of the injector spring, k_{sw} – spring constant of the injector, m_w – mass of the moving parts of the injector, p_i – pressure in the injector nozzle chamber, p_s – pressure in the injector sack, τ – time.

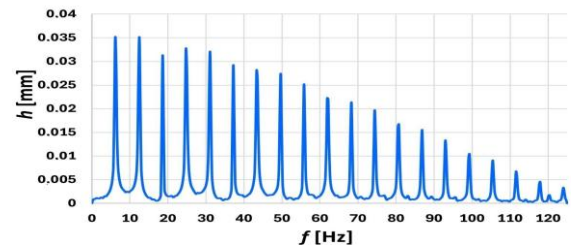
The vibrations generated and propagated during fuel injection have characteristic frequencies that are integer multiples of the rotational frequency f_o (polyharmonic signals). Functional components for diagnostic purposes were sought at these frequencies. The parameters of working and accompanying processes were used and analyzed in various domains [2].

Examples of vibration displacement measurement signals of the injector nozzle in the time domain for the unfit state (four

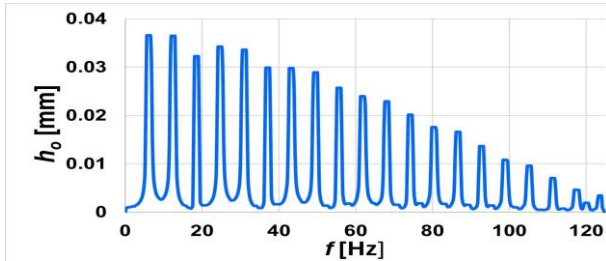
holes) are shown in Fig. 4a. The frequency domain analysis is shown in Fig. 3.4b. An envelope (Fig. 3.4c) was used to accurately determine the amplitudes of the signal's accurately spectral components. The spectrum envelope is helpful when there are slight variations in the frequencies of the individual components of an object. These are polyharmonic spectra, which include multiples of the rotational frequency in addition to the fundamental frequency (3.4d).



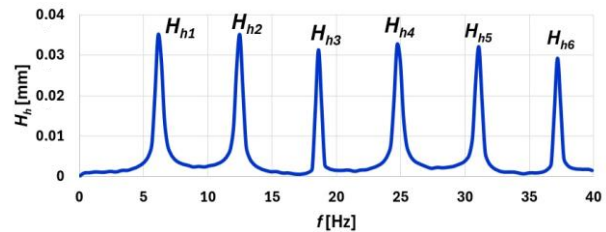
(a)



(b)



(c)



(d)

Figure 4. Example of a time waveform in a damage state (a), line spectrum (b), amplitude spectrum envelope of an analysis of the vibration displacement signal of an injector nozzle operated at 58% load (c), and view the first six amplitude components of the spectra (d).

Fig. 4b–d show that the obtained polyharmonic signals exhibit decreasing amplitudes of successive components, indicating a lower signal-to-noise ratio, and the harmonic powers, recorded as both average and root-mean-square values, also decrease. The angle of inclination of the decrease in the component amplitudes depends on the fuel rail setting, and the decrease in the component amplitudes was unambiguous.

The histogram (Fig. 5a), cumulative histogram (Fig. 5b), and values of statistical measures of the Haar wavelet approximation at level 5 (Fig. 5c) were determined for the original signal spectra of the vibration displacement of the injector nozzle (Fig. 5b). It is possible to apply the approximation reconstruction at different levels.

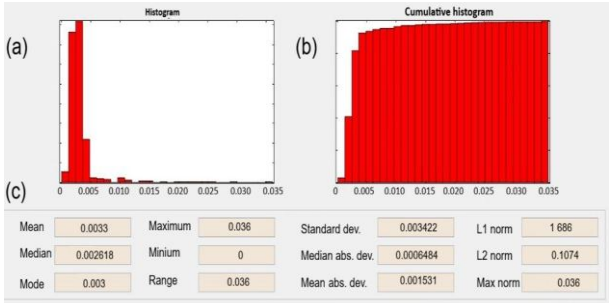


Figure 5. Results of applying the approximation reconstruction for the vibration displacement spectrum at level 5 (a), its histogram (b), and the values of statistical measures of Haar wavelet approximation (c).

Measurements of important injector state features were also carried out using various instruments for geometric measurements, mass stream of calibration oil, or airflow, an essential contribution of the authors [54]. In addition to the equipment and methods presented in Fig. 2 and described in [49,52], digital and opto-digital microscopes, a Volumetric Scanner, 3D scanners, etc., were used to measure the technical state features of the injector nozzles under study [55–60].

Fig. 6a shows the diameter of the spray hole measured using a VHX-6000 digital microscope. The other geometric features of injector nozzles, including the angle of the nozzle needle (Fig. 6b), can also be measured.

On the other hand, Table 2 presents an excerpt from the matrix containing measurement data in the frequency domain, the state feature values, and the performance parameters of the injection subsystem. Table 2 shows exemplary results of tests of the values of diagnostic parameters of the amplitudes of the envelopes of the components of the spectra from H_{h1} to H_{h6} for

various state features: values of injector opening pressure P_w , number of o spray holes i , spray hole diameters d , mean diameter clearance L_a , maximum injector nozzle stroke h_{max} , and sealing cone angle of nozzle needle α . The values of the diagnostic parameters depend on the test conditions defined by camshaft rotation speed n , fuel rod setting N , injection pump fuel feed pressure p_z , and fuel temperature at the injection pump inlet t_p . The technical state was defined as binary, with 1 indicating the ability state and 0 the disability state. The number of experimental and reference data was very large, so data and knowledge databases were built to collect them.

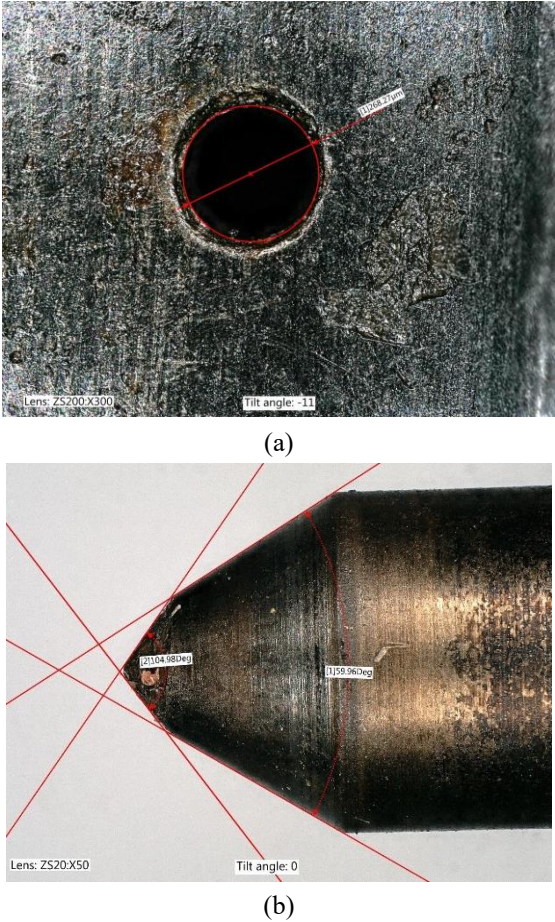


Figure 6. Measurement of the diameter of the spray hole of an injector nozzle with 300x magnification (a) and the face angle of the nozzle needle measured with magnification 50x.

It can be seen from Table 1 shows that the number of selected diagnostic parameters of the amplitudes of the first six components of the polyharmonic spectrum is considerable. In addition to a large amount of data, an averaging of 100 signal time courses was performed during the analyses. At a significance level of 0.05, there was no reason to reject the hypothesis that the means are equal for the 16- and 32-average

signals. This should safeguard against instability when evaluating the regression lines as the number of cases changes. Studies were also conducted on other diagnostic signals, such as sound pressure, fuel pressure in the injection subsystem, cylinder pressure in internal combustion engines, and vibration velocity and acceleration [5]. However, to date, the injector needle displacement parameters have proved the most

statistically promising.

An attempt was made to assess the influence of different parameters of test conditions and technical state features, i.e., explanatory variables, on diagnostic parameter values, and the selection of the most useful diagnostic symptoms was performed using multiple regression in the Statistica program.

Table 2. Experimental data matrix, a fragment of diagnostic symptoms, and technical state features.

H_{h1}	H_{h2}	H_{h3}	H_{h4}	H_{h5}	H_{h6}	P_w	i	d	L_a	h_{max}	α	n	N	p_z	t_p	B	State
[mm]	[mm]	[mm]	[mm]	[mm]	[mm]	[MPa]	–	[mm]	[mm]	[mm]	[°]	rpm	[%]	MPa	°C	$10 \frac{\text{mm}^3}{\text{inj.}}$	–
0.0298	0.0292	0.0281	0.0276	0.0267	0.0238	25	7	0.26	0.0035	0.463	60.95	375	58	0.15	20	53.0	1
0.0771	0.0736	0.0632	0.0582	0.0413	0.0352	25	7	0.26	0.0035	0.463	60.95	375	100	0.15	20	135.2	1
0.0552	0.0523	0.051	0.0415	0.041	0.0365	25	4	0.26	0.0035	0.461	61.10	375	58	0.15	20	45.5	0
0.0468	0.0449	0.0439	0.0371	0.0393	0.0347	25	5	0.26	0.0035	0.465	61.13	375	58	0.15	20	46.4	0
0.0424	0.0416	0.0407	0.0374	0.0378	0.0344	25	6	0.26	0.0035	0.465	61.03	375	58	0.15	20	49.0	0
0.0298	0.0292	0.0281	0.0276	0.0267	0.0238	25	7	0.26	0.0035	0.463	60.97	375	58	0.15	20	53.0	1
0.1408	0.1207	0.0986	0.0725	0.0399	0.0168	25	4	0.26	0.0035	0.461	61.10	375	100	0.15	20	125.1	0
0.1210	0.1082	0.0928	0.0654	0.0482	0.0244	25	5	0.26	0.0035	0.465	61.13	375	100	0.15	20	131.4	0
0.1123	0.1071	0.0814	0.0690	0.0515	0.0326	25	6	0.26	0.0035	0.465	61.03	375	100	0.15	20	132.7	0
0.0771	0.0736	0.0632	0.0582	0.0413	0.0352	25	7	0.26	0.0035	0.463	60.97	375	100	0.15	20	53.0	1

3.3. Using multiple regression to analyze experimental data

Multiple regression analysis extends the straightforward method to include multiple explanatory variables. It is more complex than linear regression regarding the best-fit model [12]. Multiple regression is recommended for non-linear models and polynomial models of second or higher degree.

In internal combustion engine research, more than two factors are usually investigated, and a relationship between a single output factor and several simultaneous input factors is sought, using multiple correlation. It assigns the values of multiple input factors to the mean values of the dependent variable according to the relationship:

$$R = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6)$$

where: y_i is the empirical value of the i -th outcome factor, $\hat{y}_i$ is the value of the i -th outcome factor calculated from the regression equation, $\bar{y}$ is the mean value of the outcome factor, and n is the number of measurements.

The closer R is to unity, the more faithful the representation of trait variability by the linear relationship of the multiple regression. The formula also determines the multiple correlation coefficient:

$$R_{y|x} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (7)$$

where: $R_{y|x}$, $R_{y|z}$, and $R_{x|z}$ – calculated linear correlation coefficients between factors; where: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. The linear correlation coefficient, measuring the strength of the relationship between variables x and z , excluding the influence of the variable y , is expressed by the following formula

$$R_{z|x,y} = \frac{R - R_{z|y} \cdot R_{y|x}}{\sqrt{(1 - R_{z|y}^2)(1 - R_{y|x}^2)}} \quad (8)$$

The partial correlation coefficient measures the correlation between a dependent variable and a specific independent variable, excluding the influence of other variables on this relationship.

The approximation of test results in previous publications mainly used fifth-degree polynomials [5,61]. Analysis using polynomial regression is sometimes used when the explanatory variable depends on only one independent variable, but the linear regression model may not be sufficiently accurate [12].

A multiple linear regression model means that the dependent variable can be represented as a linear combination of the independent variables and a random component:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (9)$$

where: Y – dependent variable, X_i – independent variables (predictors), β_i – model parameters (numerical coefficients), ε –

random component.

Model Assumptions:

- The model is linear with respect to the parameters, i.e., for all and the following equality holds:
- $$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} \quad (10)$$
- The number of observations n must be greater than or equal to the number of estimated parameters ($n \geq k + 1$).
 - None of the independent variables is a linear combination of other independent variables.
 - The random component ei has an expected value of 0 (random components, i.e., residuals, are determined as differences between the empirical values and those calculated based on the regression equation).
 - The variance of the random component ei is the same for all observations.
 - The random components are uncorrelated, meaning ei and ej are independent of each other for all i and j .
 - Each random component has a normal distribution.

Methodology for multiple regression was applied:

- 1) Analyzing scatterplots of the dependent variable Y versus the independent variables X_i to verify whether the relationship is linear.
- 2) Assuming a hypothetical linear model (Formula 9).
- 3) Estimation of model parameters (plus estimation errors).
- 4) Checking the regression analysis assumptions.
- 5) Verification of the model's content (value ranges, units, type of objects studied, data collection method, etc.).
- 6) Statistical verification of the model.
- 7) Statistical verification of the model consists of the following steps:
 - Testing the significance of the entire model.
 - Testing the significance of model parameters.
 - Assessing variable redundancy (multicollinearity).
 - Selecting appropriate predictors: strongly correlated with the dependent variable, very weakly correlated with each other.
 - Assessing goodness of fit: coefficient of determination, root mean square residual (average of the sum of error squares).
- 8) Residual analysis.

4. Selected research findings

4.1. Examples of the results of tests using multiple regression

The experiments performed collected a data resource with 6 H_{hi} explanatory variables and 11 dependent variables with range of variation (nominal values and tolerance ranges)

- injector opening pressure $P_w = 15, 20, 25$ and 30 MPa ($P_w = 25^{+5}_{-5}$ MPa);
- number of unobstructed spray holes $i = 4, 5, 6, 7, 9$ ($i = 7$ or 9);
- diameter of spray holes $d = 0.23, 0.25, 0.26, 0.28$ ($d = 0.26^{+0.010}_{-0.005}$ mm);
- average diameter clearance between the nozzle body and needle L_a ($L_{min} = 0.004, L_{max} = 0.005$ mm),
- maximum needle stroke ($h_{max} = 0.50^{0}_{-0.05}$ mm),
- sealing cone angle of nozzle needle $\alpha = 61 \pm 1^\circ$,
- camshaft speed from 150 to 450 rpm ($n = 375$ rpm $\pm 5\%$);
- fuel rail setting $N = 20, 30, 40, 50, 60, 70, 80, 90$ and 100% ($N = 58\%$ and 100%);
- injection pump fuel feed pressure $p_z = 0.1, 0.15, 0.2, 0.25$ and 0.3 MPa ($p_{zmax} = 0.15\text{--}0.4$ MPa ($p_{znom} = 0.15$ MPa));
- fuel temperature at injection pump inlet $t_p = 20, 30, 40$ and 50°C ($t_p = 20$ or 40°C).

Despite the relatively high correlation between variables H_{hi} and B ($R = 0.780$) (Fig. 7), it isn't easy to define their relationship as linear ($10 \times B$ [mm³/injection]).

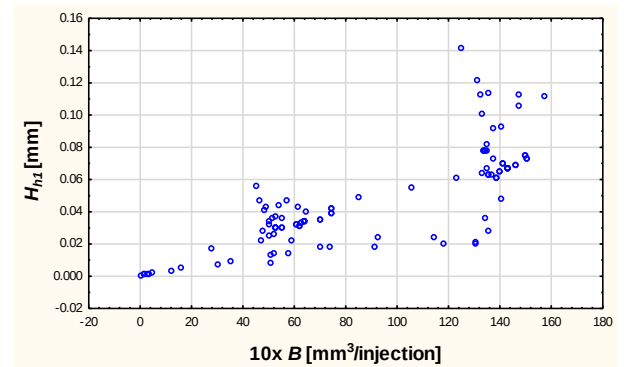


Figure 7. Scatter plot of the dependent variable H_{hi} against the independent variable B .

The correlation table between all analyzed variables (Table 3) shows some linear correlations with the dependent variable.

The most linear relationships are between $\ln(H_{hi})$ and $\log(B)$ (Fig. 7) and $\ln(H_{hi})$ and n (Fig. 8). The relationship between $\ln(H_{hi})$ and fuel rack setting N is omitted because of the strong

correlation of B with N , since the injected fuel dose B depends on the fuel rack setting N . Table 4 summarizes the regression of the dependent variable of the amplitudes of the first component of the envelope of the vibration displacement spectrum as a function of the logarithm of the injected fuel dose $\log(B)$.

Table 3. Correlations between the dependent variable H_{hl} and all analyzed independent variables.

Variable	Correlations. The labeled correlation coefficients are significant with $p < 0.05$, $N_v = 125$		
	Mean	Standard deviation	$\ln(H_{hl})$
$\ln(H_{hl})$	-3.3704	1.08430	1.000000
B	89.6280	45.42686	0.748402
$\log(B)$	1.8434	45.42686	0.923173
P_w	24.8400	1.54711	-0.09284
I	6.9680	0.59482	-0.095518
D	0.2594	0.00669	-0.067563
L_a	0.0037	0.00089	0.106563
h_{max}	0.4672	0.01760	0.105210
α	60.8527	0.34959	-0.120410
n	342.5400	72.12499	0.620388
N	74.8960	23.56522	0.767709
p_z	0.1580	0.03384	0.105512
t_p	21.9200	6.43779	0.073597

Of the independent variables analyzed, P_w , d , i , L_a , h_{max} , α (α), n , N , p_z , t_p , B , the multiple regression model showed that some variables were statistically insignificant (Table 4). Many independent variables were used in the multiple regression, but some were insignificant. During injector testing in the manufacturing and maintenance phase, the goal is to measure as few control quantities as possible while minimizing production time and costs, and ensuring high-quality production and maintenance. Therefore, selecting a global measure, such as fuel injection dose or fuel mass stream, combined with other statistically selected features, is a solution to expectations.

Table 4. Summary of the regression of the dependent variable $\ln(H_{hl})$ of the amplitudes of the first component of the envelope of the vibration displacement spectrum as a function of the logarithm of the injected fuel dose $\log(B)$.

$N_v = 125$	Regression summary of dependent variable: $\ln(H_{hl})$, $R = 0.923173$, $R^2 = 0.852249$. Corrected $R^2 = 0.851048$, $F(1.123) = 709.48$ $p < 0.0000$, standard error of estimation 0.41848					
	b^*	Error std. of b^*	b	Error std. of b	$t(123)$	p
y-intercept			-7.74549	0.168465	-45.9767	0.00
$\log(B)$	0.923173	0.034659	2.37343	0.089106	26.6361	0.00

b^* – standardized coefficient, b – regression coefficients, $t(n)$ – test value with n degrees of freedom, p – test probability, R – correlation coefficient, R^2 – coefficient of determination.

Only the predictors $\log(B)$, n , and B have a significant statistical relationship with the dependent variable H_{hl} (the amplitude of the envelope of the first components of the needle

displacement spectra). The significance analysis of the multiple regression model coefficients partially confirmed this.

Fig. 8 shows the dependence of the scatter of $\ln(H_{hl})$ on the injected fuel dose B . It can be seen from Fig. 8 that the scatter of the results is not very large. A linearized regression model was used, in which various transformations reduce the nonlinear model to a linear problem. The authors achieved this using logarithmic transformations.

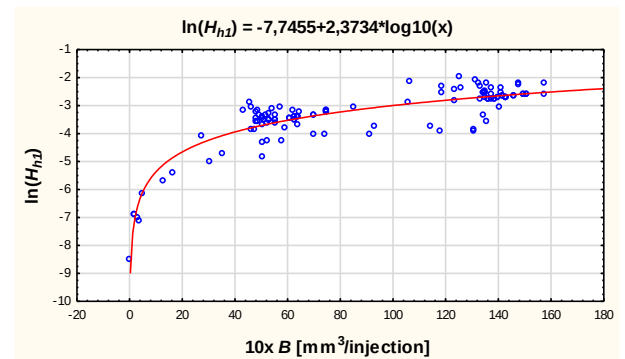


Fig. 8. Scatter plot of $\ln(H_{hl})$ against the injected fuel dose B .

Fig. 10 shows that the attempt to linearize the results was successful, and they dissipated in the different technical states of the injector nozzles due to simulated damage and random wear. Similarly, the scatter plot of $\ln(H_{hl})$ from camshaft speed n is shown (Fig. 9), where the high concentration of measurement points at $n = 375$ rpm is due to this being the central value of camshaft speed in the experiments conducted.

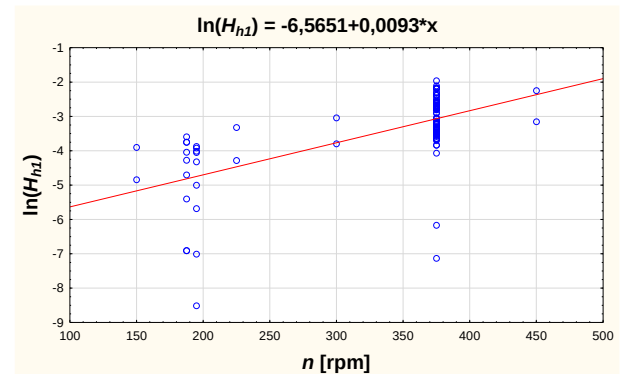


Fig. 9. Dependence of the natural logarithm of the dependent variable H_{hl} on camshaft speed n .

Considering only significant correlations for the dependence of H_{hl} on two variables, $\log(B)$ and camshaft speed n , the data in Table 5 were obtained. It can be seen from Table 5 that there is a very good fit, as the coefficient of determination is $R^2 = 0.96825407$. Furthermore, a multiple regression analysis was performed, first for all cases and then for the cases that were not removed. The results turned out to be very similar.

Multiple linear regression assumes that the relationship between variables is linear. To demonstrate the validity of this assumption, two-dimensional scatter plots of the experimental variables studied were analyzed. If the relationship was nonlinear, the variables were transformed, or nonlinear components were used. In a sense, the straight line fits the set of experimental points shown in Fig. 9 and Fig. 10.

Table 5 lists the values of the regression equation coefficients, their standard errors, and the test statistic for their statistical significance. A test probability value p of approximately 0 indicates that the equation coefficients are statistically significant. A very good fit was obtained in Table 5, with $R^2 = 0.968$. Table 5 presents the results of multiple regression modeling for the two significant factors: $\log(B)$ and

Table 5. Summary of the dependent variable regression with only significant correlations for the dependence of H_{hl} on two variables: $\log(B)$ and camshaft speed n for $N_v = 125$ and $N_v = 118$ injectors.

Regression summary of the dependent variable: $\ln(H_{hl})$, $R = 0.972117$, $R^2 = 0.945012$ Corrected $R^2 = 0.944110$, $F(2.122) = 1048.3$ $p < 0.0000$, standard error of estimation 0.25634						
$N_v = 125$	b^*	Error std. of b^*	B	Error std. of b	$t(122)$	p
y-intercept			- 8.86441	0.129353	- 68.5287	0.00
$\log(B)$	0.803827	0.022802	2.06659	0.058623	35.2524	0.00
n	0.327118	0.022802	0.00492	0.000343	14.3460	0.00
Regression summary of the dependent variable: $\ln(H_{hl})$, $R = 0.98399902$, $R^2 = 0.96825407$. Corrected $R^2 = 0.96770197$, $F(2.115) = 1753.8$ $p < 0.0000$, standard error of estimation 0.18083						
$N_v = 118$	b^*	Error std. of b^*	B	Error std. of b	$t(115)$	p
y-intercept			- 8.68515	0.093425	- 92.9644	0.00
$\log(B)$	0.761636	0.018298	1.91378	0.045977	41.6247	0.00
n	0.380915	0.018298	0.00519	0.000249	20.8176	0.00

b^* – standardized coefficient, b – regression coefficients, $t(n)$ – test value with n degrees of freedom, p – test probability, R – correlation coefficient, R^2 – coefficient of determination.

Table 6. Contribution of variables to the variance of H_{hl} .

Currently, there are variables in the equation; $\ln(H_{hl})$, sheet new logarithmic data H_{hl} , $v3^*v3$ – square of semipartial correlation (SSC)								
Variable	b^*	Partial correlation	Semi-partial correlation	Tolerance	R^2	$t(115)$	p	$v3^*v3$
$\log(B)$	0.761636	0.968379	0.691586	0.824512	0.175488	41.62469	0.00	0.478291
n	0.380915	0.888982	0.345881	0.824512	0.175488	20.81762	0.00	0.1196336

b^* – standardized coefficient, b – regression coefficients, $t(n)$ – test value with n degrees of freedom, p – test probability, R – correlation coefficient, R^2 – coefficient of determination, SSC – square of semipartial correlation.

The contribution of variables to the variance of H_{hl} is shown in Table 6.

In multivariate regression, the residuals are assumed to be normally distributed. Therefore, it is good to examine the distributions of the main variables under study by developing histograms and normality plots for the residuals before formulating conclusions. This also comes out when checking one of the basic regression assumptions about the normality of the residuals. The assumption of normality of the residuals is satisfied. The residuals follow be normally distributed (Fig. 11a

n .

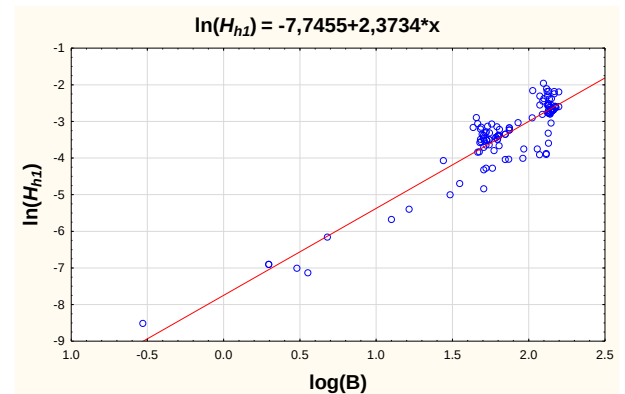


Fig. 10. Scatter of the natural logarithm of the amplitude of the first $\ln(H_{hl})$ component against $\log(B)$.

and 11b). Fig. 11 does not show outliers because cases 3, 4, 7, 8, 45, 46, and 122 were omitted.

Table 7 shows the standardized residuals of $\ln(H_{hl})$, and Table 8 summarizes the regression of the dependent variable $\ln(H_{hl})$ for the six predictors.

For the case of no. 45, the values are outliers because, at the maximum fuel rod setting (and therefore a hefty dose of injected fuel), the camshaft rotational speed slowed, and thus the amplitudes of the components of the displacement spectra of the injector needle were large.

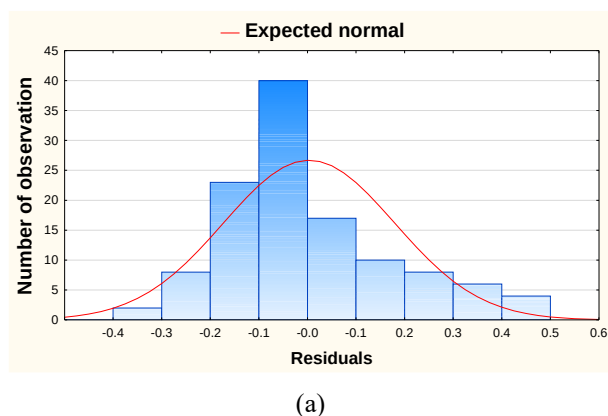
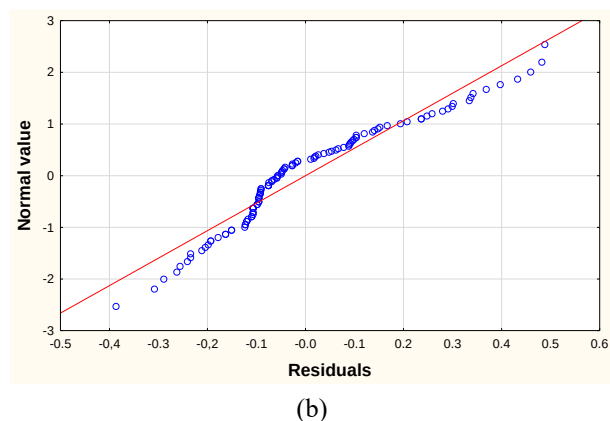


Fig. 11. Distribution of raw residuals (a) and normality plot of residuals (b).

Table 7. Standardized residuals of $\ln(H_{h1})$.

Standardized residuals								Standardized residuals $\ln(H_{hi})$					
Case	-5	-4	-3	± 2	3	4	5	Observed value	Predicted value	Residual	Predicted standard	Standard rest	Standard error
3	.	.	.		*	.	.	-2.89679	-3.59381	0.69702	-0.21197	2.71913	0.029254
4	.	.	.		.	.	.	-3.06187	-3.57623	0.51436	-0.19529	2.00657	0.029003
7	.	.	.		.	.	.	-1.96042	-2.68607	0.72566	0.64921	2.83084	0.027389
8	.	.	.		*	.	.	-2.11197	-2.64197	0.53001	0.69104	2.06761	0.027907
45	.*	.	.		.	.	.	-7.13090	-5.87766	-1.25324	-2.37867	-4.88899	0.083588
46	.	.	.		*	.	.	-6.16582	-5.61408	-0.55173	-2.12862	-2.15235	0.076491
122	.	.	.		*	.	.	-2.16456	-2.83071	0.66615	0.51199	2.59869	0.026042
Minimum	.*	.	.		.	.	.	-7.13090	-5.87766	-1.25324	-2.37867	-4.88899	0.026042
Maximum	.	.	.		*	.	.	-1.96042	-2.64197	0.72566	0.69104	2.83084	0.083588
Average	.	.	.		*	.	.	-3.64176	-3.83151	0.18975	-0.43747	0.74021	0.042811
Median	.	.	.		.	.	.	-2.89679	-3.57623	0.53001	-0.19529	2.06761	0.029003

* – at what distance from the expected value is the rest.

Finally, the dependence of the dependent variable H_{hl} on the two most important properties of the technical state is presented in the relationship:

$$H_{h1} = 0.00017 \exp(1.91378 \log(B) + 0.00519 n \pm 0.18083) \quad (11)$$

For the six predictors, the regression results summary is shown in Table 8. Table 8 presents the results of multiple regression analysis using the six most important factors. The difference in goodness of fit between the models was small: $R^2 = 0.968$ for the first model and $R^2 = 0.979$ for the second. Given that the linearity of the relationship and appropriate assumptions about residuals were maintained in the first case, only the model with two independent variables is statistically significant.

The effect of the independent variables P_w , i , d , and h_{max} on the variation of H_{hl} is negligible and less than 1% (Table 9).

Slightly worse results were obtained for the variable for the envelope amplitudes of the second component of the vibration displacement spectrum of the injector needle H_{h2} and further components up to H_{h6} . Due to the limited size of the article, only the relationships of the dependent variables H_{h2} to H_{h6} on the two most essential properties of the technical state are presented:

$$H_{h2} = 0.00018 \exp(1.89949 \log(B) + 0.00496 n \pm 0.25676) \quad (12)$$

$$H_{h3} = 0.00018 \exp(1.75003 \log(B) + 0.00552 n \pm 0.20590) \quad (13)$$

$$H_{h4} = 0.00020 \exp(1.64957 \log(B) + 0.00554 n \pm 0.23328) \quad (14)$$

$$H_{h5} = 0.00022 \exp(1.40187 \log(B) + 0.00600 n \pm 0.25713) \quad (15)$$

$$H_{h6} = 0.00028 \exp(1.18033 \log(B) + 0.00608 n \pm 0.32232) \quad (16)$$

The camshaft rotational speed appears in the formula for the frequency of vibrations generated in the injection subsystem (2). In equations (9) to (14), the influence of the values of the coefficients at $\log(B)$ and n is significant.

The degree of association between the explanatory variables and the dependent variable H_{hi} is expressed by the value of the correlation coefficient R , which is the square root of R^2 . In a multiple regression, R takes values between -1 , 0 , and 1 . Associations between variables can be positive or negative (e.g., as P_w increases, H_{hi} decreases). The criteria for dividing the strength of the relationship between two characteristics according to the value of the correlation coefficient are as follows [60,62]:

$0.8 \leq R \leq 1.0$	$\Rightarrow$	very strong relationship,
$0.6 \leq R < 0.8$	$\Rightarrow$	strong relationship,

$0.4 \leq R < 0.6$ $\Rightarrow$ moderate relationship, $0.0 \leq R < 0.2$ $\Rightarrow$ very weak relationship.
 $0.2 \leq R < 0.4$ $\Rightarrow$ weak relationship,

Table 8. Regression summary of the six predictors' dependent variable $\ln(H_{hl})$.

Dependent variable regression summary: $\ln(H_{hl})$, $R = 0.98922273$, $R^2 = 0.97856162$. Corrected $R^2 = 0.97740279$, $F(6.111) = 844.44$ $p < 0.0000$, standard error of estimation 0.15126						
$N_v = 118$	b^*	Error std. b^*	B	Error std. b	$t(111)$	P
y-intercept			-4.3159	0.843627	-5.11587	0.000001
$\log(B)$	0.761533	0.015309	1.9135	0.38466	49.74552	0.000000
n	0.372257	0.015435	0.0051	0.000210	24.11834	0.000000
P_w	-0.058310	0.013926	-0.0368	-0.008801	-4.18718	0.000057
I	-0.063615	0.016287	-0.1646	0.042140	-3.90586	0.000162
D	-0.092307	0.016484	-13.4906	2.409142	-5.59975	0.000000
h_{max}	0.047895	0.014422	2.6647	0.802367	3.32100	0.001215

b^* – standardized coefficient, b – regression coefficients, $t(n)$ – test value with n degrees of freedom, p – test probability, R – correlation coefficient, R^2 – coefficient of determination.

Table 9. Partial correlation indices for dependent variable H_{hl} , $\log(B)$, and five other variables.

Currently, there are variables in the equation; $\ln(H_{hl})$, sheet new logarithmic data H_{hl} , $v3*v3$ – SSC								
Variable	b^*	Partial correlation	Semipartial correlation	Tolerance	R^2	$t(111)$	p	New variable = $v3*v3$
$\log(B)$	0.761533	0.978300	0.691335	0.824137	0.175863	49.74552	0.000000	0.47794385
n	0.372257	0.916382	0.335183	0.810732	0.189268	24.11834	0.000000	0.11234762
P_w	0.058310	-0.369330	-0.058191	0.995907	0.004093	-4.18718	0.000057	0.00338619
i	-0.063615	-0.347609	-0.054281	0.728083	0.271917	-3.90586	0.000162	0.00294647
d	-0.092307	-0.469331	-0.077822	0.710787	0.289213	-5.59975	0.000000	0.00294647
h_{max}	0.047895	0.300633	0.046153	0.928596	0.071404	3.32100	0.001215	0.00213013

b^* – standardized coefficient, b – regression coefficients, $t(n)$ – test value with n degrees of freedom, p – test probability, R – correlation coefficient, R^2 – coefficient of determination, SSC – square of semipartial correlation.

The model's goodness of fit is very high, often $R^2 > 0.9$. The F -statistic, which verifies the model's linearity, is very high ($F(6, 102) = 130$, $p = 0$). All coefficients of the multiple regression equation are statistically significant.

The test for regression significance determines whether a linear relationship exists between the dependent variable and a subset of the regressor variables. The appropriate hypotheses are $H_0: b_1 = b_2 = \dots = b_k = 0$ (k – number of regressor variables), $H_1: b_j \neq 0$ for at least one j .

Rejection of H_0 implies that at least one of the regressor variables contributes significantly to the model. In our research, the F -statistic is $F(6, 102) = 130$. Critical value for significance level 0.05, with degrees of freedom 6 and 102 is equal to 2.19. This means that the value of the test statistic is significantly greater than the critical value in the significance test (P -value is considerably smaller than $\alpha = 0.05$). So, we reject the null hypothesis and conclude that the variable $\ln(H_{hl})$ (natural logarithm of the amplitude of the envelope of the first components of the needle displacement spectra) is linearly related to two regressor variables: $\log(B)$ and n .

The coefficient of determination, R^2 , is a measure of the amount of variation in the dependent variable explained by the regressor variables. The coefficient of determination (R^2) for the

regression model fit to $\ln(H_{hl})$ is greater than 0.9 in our research ($R^2 > 0.9$). That is, about 90% of the variability in the amplitude of the envelope of the needle displacement spectral first components is explained by the two regressor variables, $\log(B)$ and n , used in [63].

In equations (9–14), each variable from $\log(B)$ to n is correlated with the variable H_{hl} after considering the effect of the other independent variables included, which is called partial correlation [60]. The regression coefficients next to the variables show the contributions of each independent variable to the prediction of the dependent variable H_{hl} .

The goodness of fit of the best multivariate regression model for individual H_{hi} is as follows (Fig. 12). Fig. 12 shows that the best dependent variable is H_{hl} , followed by H_{h3} , and the weakest is H_{h6} .

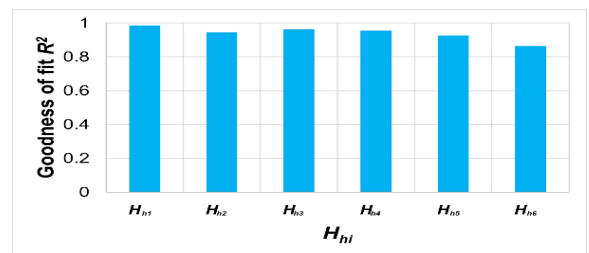


Figure 12. Plot of the goodness-of-fit coefficients of the regression model for the dependent variable H_{hi} .

From the determined frequencies of hydrodynamic and mechanical phenomena, it follows that the first component is fundamental ($i = 1$). This was confirmed by statistical analyzes, which is why the usefulness of H_{h6} is the lowest (Table 1).

In some analyzes, outliers can strongly affect the results. The most common cases are 45, 46, 51, and 123. The outlier points distort the results by attracting or moving the regression line in a particular direction, causing the values of the coefficients to change.

In each case of H_{hi} , one can try to study the outliers and omit other elements to see what will be more favourable. For example, in H_{h1} , there are no outliers, and the normality of the residuals doesn't work out either.

As H_{hi} changes, the contribution of $\log(B)$ to explaining the variance of H_{hi} decreases. The contribution of speed n to explaining the variance of H_{hi} increases (Fig. 13). The pure (without the contribution of the other variables) contribution of a predictor to the variance of the dependent variable is determined by the semi-partial correlation, and its square can be interpreted as the specific contribution of a given predictor to the variance of H_{hi} (in %).

$\log(B)$ and n are the two variables that most affect the variance of H_{hi} . This means that they are the most important variables in injector diagnostics.

The analyses presented here demonstrate the existence of *relationships* among the variables, but will it be possible to demonstrate the causal relationships of these relationships, i.e., that the selected diagnostic symptom analysis will show a significant effect of the technical state features? The causal inference analysis procedure begins by selecting the “**Multiple Regression**” analysis in the relevant program.

Fig. 14a show the effect of camshaft rotational speed, and Fig. 14b shows the effect of the number of spray holes on the envelope amplitudes of the nozzle needle displacement spectra. The figure shows that increasing rotational speed increases the amplitudes of the displacement-spectrum components of the nozzle needle, whereas increasing the number of orifices decreases them.

4.2. Procedure for diagnostic testing of injectors

During injector testing, in manufacturing and maintenance, the goal is to measure as few control quantities as possible, while

minimizing production time and costs, while ensuring high quality of production and maintenance. Therefore, selecting a global measure, such as fuel injection dose or fuel mass flow rate, combined with other statistically selected characteristics, is a solution to expectations.

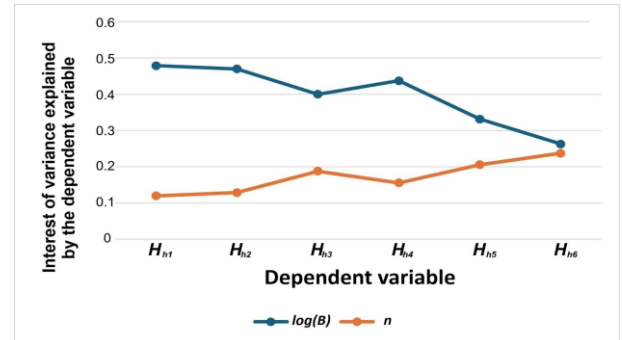
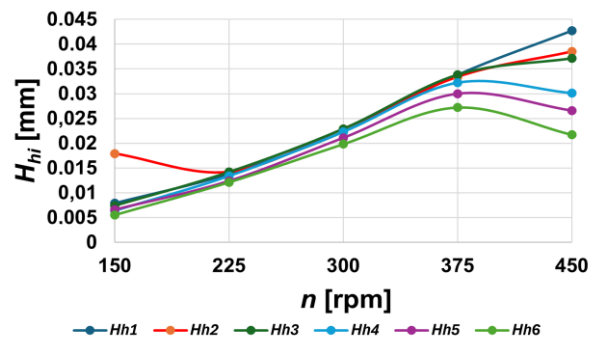
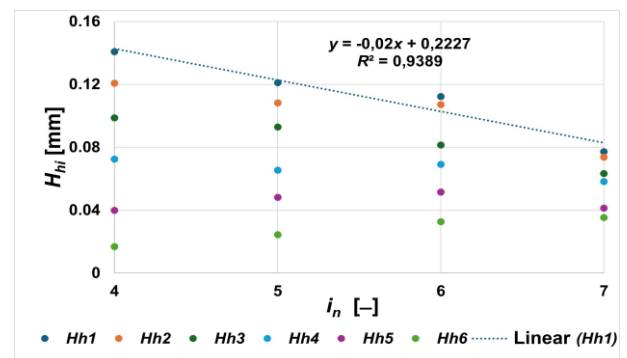


Fig. 13. Contribution of $\log(B)$ and n variables in explaining H_{hi} variance.



(a)



(b)

Fig. 14. Effect of camshaft rotational speed at $N = 58\%$ (a) and effect of the number of open spray holes on the envelope amplitudes of the displacement spectrum components of the nozzle needle at $N = 100\%$ (b).

For injectors in operation, the value of camshaft rotational speed is $n = 375$ rpm, and this variable cannot be used in multiple regression. A univariate all n regression was used for comparison: the relationship between $\ln(H_{h1})$ and $\log(B)$. This

indicates that the correlation coefficients are similar, with $R = 0.925$ for simulated injector states and $R = 0.906$ for operated injectors. Similarly, the coefficients of determination differ slightly: $R^2 = 0.855$ for simulated injector states and $R^2 = 0.820$ for operated injectors. The fuel dose B can be calculated by integrating the fuel mass flux q_{mi} flowing through the number of spray holes using the following equation [55]:

$$B = \int_{ea}^{sa} q_{mi} = \alpha_n \epsilon i \frac{\pi d_i^2}{4} \int_{ea}^{sa} \sqrt{2(p_i - p_c) \rho_{p,t}} \quad (15)$$

where: d_i – diameter of the spray hole in temperature t , α_n – flow number, $\Delta p = p_i - p_c$ – pressure difference in the injector nozzle space p_i , and in the combustion chamber p_c , ϵ – expansion number, $\rho_{p,t}$ – density of the flowing fuel, determined for conditions immediately before the spray hole, at pressure p and temperature t , sa – fuel injection start angle, ea – fuel injection end angle.

The evaluation of diagnostic parameters was based on the relationship between variables, as measured by the correlation coefficient. Cumulants expressed diagnostic parameters and symptoms measured at different measurement points:

$$K_{ui}(\theta) = \sqrt[n]{P_{i1}(\theta)P_{i2}(\theta) \dots P_{in}(\theta)} \quad (16)$$

where: $K_{ui}(\theta)$ – the cumulant of the i -th signal parameter, $P_{in}(\theta)$ – the value of the i -th parameter at the n -th measurement point. The cumulant is the geometric mean.

Table 10 shows the signal suitability analysis of the injection subsystem pressure p_i , cylinder pressure p_c , nozzle needle vibration displacement h , velocity v , and vibration acceleration a , based on the correlation coefficient R , sensitivity coefficient w_s , coefficient of variation w_v , and i -cumulants K_{ui} .

After preliminary research, the most promising diagnostic symptoms were found to be the amplitudes of the envelope components in the spectra of the injector nozzle needle displacement signal.

Based on the results obtained from the active experiment and operational tests of AL20/24 engine injectors, the following procedure for diagnosing injectors is proposed:

- 1) For each injector, after manufacture, control tests should be carried out for fuel rod settings $N = 58\%$ and $N = 100\%$, determining the amplitude-frequency spectra of needle displacement in the 0...100 Hz frequency band.
- 2) The tests are performed on a test bench (test station,

Fig. 2) at a camshaft rotational speed of 375 rpm.

- 3) The injectors should be prepared for inspection by installing a lifting mandrel and an adjustment screw suitable for measurements, which should be replaced with the original ones after the test. The preparation consists of leading the axial extension of the lifting mandrel outside the outline of the adjustment screw (Fig. 1b).
- 4) After each measurement, it is necessary to check whether the values of the amplitudes of the injector spectrum envelope components fall within the ranges of the suitability state interval between the limit values:
- 5) a) $N = 58\%$ $0.0268 < H_{h2} < 0.0329$ mm,
- 6) $0.0254 < H_{h3} < 0.0312$ mm,
- 7) b) $N = 100\%$ $0.0728 < H_{h3} < 0.0868$ mm.

Table 10. Matrix for selecting the most useful diagnostic signals and their symptoms in different areas: 1 – useful, R – correlation coefficient, w_s – sensitivity coefficient, w_v – coefficient of variation, K_u – cumulant.

Analysis domain	Signal/diagnostic parameter				
	P_i	P_c	h	v	a
Time	1	1	1	1	1
Amplitude	1	1	1	1	1
Frequency	1	1	1	1	1
Waveform	1	1	1	1	1
...					
R	0.972	0.994	0.990	0.978	0.990
w_s	1.863	0.588	1.875	1.849	0.575
w_v	0.308	0.364	0.698	0.573	0.219
K_u	1.346	0.765	1.362	1.344	0.754

If the measurement results fall within the specified ranges, the injector being tested is suitable. However, if the value of at least one of the three components envelope amplitudes is outside the ranges specified for the suitability state, then a time domain analysis should be performed in the frequency band up to 2.5 kHz at fuel rack settings $N = 58\%$ and $N = 100\%$. To identify damage, artificial intelligence techniques should be used to recognize damage types and signal images [2,52].

An empirical comparison of multiple linear regression and artificial neural network for concrete dam deformation modeling. ANN was used and its results were compared with the multiple linear regression (MLR) results [7]. The same variables as in the regression analysis were selected as significant predictors, namely (Table 11): P_w , d , i , h_{max} , n , B .

The MLR and ANN results were found to be strongly correlated ($R \geq 0.81$). For this purpose, 30 different input data

sets were selected, and H_{hl} values were determined using both the MLR and ANN methods.

Table 11. For the H_{hl} variable, the generated neural networks had very good parameters.

Network name	Quality (learning)	Quality (testing)	Quality (validation)
MLP 6-10-1	0.979309	0.939437	0.992581
MLP 6-5-1	0.989836	0.962242	0.992965
MLP 6-3-1	0.987982	0.958337	0.993809
MLP 6-3-1	0.989698	0.960536	0.993168
MLP 6-3-1	0.989548	0.949544	0.992788

It was found that useful symptoms for assessing the technical state of injectors (amplitudes of the envelope of the needle displacement spectrum components) are very closely related to changes in the injected fuel doses.

5. Conclusion

The results met the research objectives and expectations to a large extent. The most useful diagnostic symptoms of the amplitudes of the envelopes of the signal displacement spectra of the injector needle vibration signals of the conventional injectors component in the frequency domain of the first and third harmonics, with the relevant features of the technical state and test conditions, were selected. Few studies, in addition to the authors, have examined relationships among diagnostic symptoms, state features, and test-condition parameters.

The study showed that the determination coefficient values were higher for the simulated single-fault state than for nozzles operated under random wear. The relevance of the individual condition features was investigated by assessing the model's

best fit in the dependencies of the influence of the independent variables n , P_w , i , h_{max} , and $\log(B)$ on the diagnostic parameters. The technical state features and condition control values of injection processes of low relevance were rejected in the analyses. The strongest linear correlations are between the natural logarithm of the first component of the spectral envelope and the logarithm of the injected fuel dose, and the fuel rod setting, which linearizes the relationship. Only the predictor's logarithm of fuel dose, fuel dose, and camshaft speed are statistically significant predictors of the dependent variables.

Therefore, when inspecting the technical state of injectors during production and operation, the injected fuel dose should be checked at a precisely controlled rotational speed (injection frequency). The selected state features should be checked after production or maintenance. The injected fuel dose, fuel volume, or fuel mass stream can be dynamically measured using a fuel pressure sensor in the injection subsystem under operation conditions. The method described for diagnosing fuel injectors and measuring fuel dosage during quality control should be applied in industry.

Useful measured values were selected, and the method was validated by locating damage to components of a real internal combustion engine and for in-operation nozzle tests. The following other diagnostic signals were also examined, such as fuel pressure, velocity and vibration acceleration, as well as modern diagnostic symptoms in different analysis domains.

Acknowledgment

This research was sponsored by the Ministry of Science and Higher Education of Poland. Award Mathematical modeling of transportation system parameters Number: 1/S/IMFICH/26.

References

1. Monieta J. Diagnosing marine piston engines driving generators at different operational loads. *Journal of Marine Science and Engineering* 2021; 9(2): 132. <https://doi.org/10.3390/jmse9020132>.
2. Monieta J. Diagnosing cracks in the injector nozzles of marine internal combustion engines during operation using vibration symptoms. *Applied Sciences* 2023; 13(17): 9599. <https://doi.org/10.3390/app13179599>.
3. Bronsztejn IN, Siemiendajew KA, Musiol G, Mühlig A. *Modern compendium of mathematics*, 1st ed. PWN, Warsaw; 2009.
4. Korzyński M. *Methodology of experiment*, 2nd ed. WNT, Warsaw; 2013.
5. Monieta J., Selection of diagnostic symptoms and injection subsystems of marine reciprocating internal combustion engines. *Applied Sciences* 2019; 9(8): 1540. doi:10.3390/app9081540.
6. Li ZX, Yan X, Guo Z, Liu P, Chengqing Y., Z. Peng, A new intelligent fusion method of multi-dimensional sensors and its application to tribo-system fault diagnosis of marine diesel engines. *Tribology Letters* 2012; 47(1): 1–15. doi:10.1007/s11249-012-9948-1.
7. Li MJ, Wang JX. An empirical comparison of multiple linear regression and artificial neural network for concrete dam deformation

- modelling. *Mathematical Problems in Engineering* 2019; 2019(1): 7620948. doi:10.1155/2019/7620948.
8. Kowalski J. Concept of the multidimensional diagnostic tool based on exhaust gas composition for marine engines. *Applied Energy* 2015; 150: 1-8. doi:10.1016/j.apenergy.2015.04.013.
9. Kowalski J. The model of combustion process in the marine 4-stroke engine for exhaust gas composition assessment. *Combustion Engines* 2016; 165(2): 60-69. doi:10.19206/CE-2016-208.
10. Kowalski J, Krawczyk B, Woźniak M. Fault diagnosis of marine 4-stroke diesel engines using a one-vs-one extreme learning ensemble. *Engineering Applications of Artificial Intelligence* 2017; 57: 134-141. doi:10.1016/j.engappai.2016.10.015.
11. Żółtowski B, Łukasiewicz M, Kałaczyński T. *Information technology in machine condition testing*. University Publishing House of the University of Technology and Life Sciences in Bydgoszcz, Bydgoszcz; 2012.
12. Syska J. *Modern methods of computer-assisted regression analysis*, University of Silesia, Katowice; 2014.
13. Polanski Z. *Planning experience in creating quality*, Statsoft, Kraków, Poland; 1999: 37-46.
14. Eyvazian M, Noorossana R, Saghaei A, Amiri A. Phase II monitoring of multivariate multiple linear regression profiles. *Quality and Reliability Engineering International* 2011; 27(3): 281-296. <https://doi.org/10.1002/qre.1119>.
15. Ran Q, Yedong P, Du W, Du D, Peng X. Fault detection of diesel engine air and after-treatment systems with high-dimensional data: a novel Fault-relevant feature selection method. *Processes* 2021; 9(2): 259. <https://doi.org/10.3390/pr9020259>.
16. Lin JG, Zhuang QY, Huang C. Fuzzy statistical analysis of multiple regression with crisp and fuzzy covariates and applications in analyzing economic data of China. *Computational Economics* 2012; 39(1): 29-49, doi:10.1007/s10614-010-9223-1.
17. Ćwikła G, Paprocka I. Condition-based failure-free time estimation of a pump. *Sensors* 2023; 23(4): 1785. <https://doi.org/10.3390/s23041785>.
18. Kościelny JM. *Diagnostics of automated industrial processes*, Academic Publishing House: EXIT, Warsaw; 2001.
19. Wang X, Liu C, Bi F, Bi X, Shao K, Fault diagnosis of diesel engine based on adaptive wavelet packets and EEMD-fractal dimension. *Mechanical Systems and Signal Processing* 2013; 41(1): 581-597. doi:10.1016/j.ymssp.2013.07.009.
20. Wang XJ, Cai YP, Lin XZ. Diesel engine PT pump fault diagnosis based on the characteristics of its fuel pressure. In: *2013 International Conference on Applied Computing, Computer Science. IERI Procedia* 2014; 7: 84-89 doi:10.1016/j.ieri.2014.08.014.
21. Mahmoud MA, Phase I analysis of multiple linear regression profiles. *Communications in Statistics – Simulation and Computation* 2008; 37(10), 210; 2106–2130. doi:10.1080/03610910802305017.
22. Galindo J, Climent H, Pla B, Jimenez VD. Correlations for Wiebe function parameters for combustion simulation in two-stroke small engines. *Applied Thermal Engineering* 2011; 31(6–7): 1190-1199. doi:10.1016/j.applthermaleng.2010.12.020.
23. Sheikhi A, Arad F, Mesiar R. A heteroscedasticity diagnostic of a regression analysis with copula dependent random variables. *Brazilian Journal of Probability and Statistics* 2022; 36(2): 408-419. doi:10.1214/22-BJPS532.
24. Adamkiewicz A. Diagnosing marine gas turbine engine with the application of statistical methods. *Diagnostyka* 2004; 30(1): 5-8.
25. Rudnicki J, Zadrąg R. Problems of modelling toxic compounds emitted by a marine internal combustion engine in unsteady states. *Polish Maritime Research* 2014; 21(4): 57-65. doi:10.2478/pomr-2014-0042.
26. Girtler J, Kuszmidler S, Plewiński L. *Selected issues of marine vessel operation in terms of shipping safety*, Maritime University of Szczecin, Szczecin; 2003.
27. Zhu G, Huang L, Yin J, Gai W, Wei L. Multiple faults diagnosis for ocean-going marine diesel engines based on different neural network algorithms. *Science Progress* 2023; 106(4): 1-24. doi:10.1177/00368504231212765.
28. van der Ploeg C, Alirezai M, Wouw N, Esfahani PM. Multiple faults estimation in dynamical systems: tractable design and performance bounds. *IEEE Transactions on Automatic Control* 2022; 67(1): 4916-4923. <https://doi.org/10.48550/arkiv.2011.13730>.
29. Lim Y, Parkv J, Kang S. A method for the fast diagnosis of multiple defects using an efficient candidate selection algorithm. *IEICE Electronics Express* 2012; 9: 834-839. <https://doi.org/10.1587/elex.9.834>.
30. Antory D, Kruger U, Irwin G, McCullough G., Fault diagnosis in internal combustion engines using non-linear multivariate statistics. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 2005; 219: 243-258. doi:10.1243/095965105x9614.
31. Niu X, Yang C, Wang H, Wang Y. Investigation of ANN and SVM based on limited samples for performance and emissions prediction of

- a CRDI-assisted marine diesel engine. *Applied Thermal Engineering* 2017; 111: 1353-1364. doi:10.1016/j.applthermaleng.2016.10.042.
32. Lv YQ, Yang XT, Li YF, Liu JL, Li SJ. Fault detection and diagnosis of marine diesel engines: A systematic review. *Ocean Engineering* 2024; 294: 116798. doi:10.1016/j.oceaneng.2024.116798.
 33. Wang JX. A multivariate statistics-based approach for detecting diesel engine faults with weak signatures. *Energies* 2020; 13(4): 873. doi:10.3390/en13040873.
 34. Chalecki D. The algorithm of fault signatures minimization. *Diagnostyka* 2006; 3(39): 297-300.
 35. Cempel C. Technical diagnostics – a synthetic view. Machine condition monitoring – historic overview. *Diagnostics and Structural Health Monitoring* 2011; 4(60): 55-67.
 36. Zhang K, Ding SX, Shardt YAW, Chen ZW, Peng KX. Assessment of T2 and Q statistics for detecting additive and multiplicative faults in multivariate statistical process monitoring. *Journal of the Franklin Institute* 2017; 354(2): 668-688. doi:10.1016/j.jfranklin.2016.10.033.
 37. Taylor JR. *An Introduction to error analysis: The study of uncertainties in physical measurements*, Sausalito, 2nd ed. CA, USA: University Science Books, California; 1997.
 38. Jafari M et al. Multivariate analysis of performance and emission parameters in a diesel engine using biodiesel and oxygenated additive. *Energy Conversion and Management* 2019; 201: 112183. doi:10.1016/j.enconman.2019.112183.
 39. Sehili Y, Cerdoun M, Loubar K, Tarabet L. Dual fuel engine injector temperature monitoring: an innovative thermal analysis approach. *Applied Thermal Engineering* 2024; 249: 123370, <https://doi.org/10.1016/j.applthermaleng.2024.123370>.
 40. Abdullayev P, Pashayev A, Askerov D, Sadiqov R. Aircraft gas turbine engines' technical condition identification system. *International Journal of Aerospace and Mechanical Engineering* 2005; 8: 288-293. <https://doi.org/10.1115/gt2005-69072>.
 41. Patil C, Theotokatos G, Wu Y, Lyons T. Investigation of logarithmic signatures for feature extraction and application to marine engine fault diagnosis. *Engineering Applications of Artificial Intelligence* 2024; 138(12): 1-13. doi:10.1016/j.engappai.2024.109299.
 42. Beijer A. A Method and a system for diagnosing injection systems of self-ignition engines, particularly marine engines. EU Patent 2013; 13460015.4-1606.
 43. Han J, Mu X, Liu Y, Wang X, Li W, Bai C, Zhang H. A critical review of on-line oil wear debris particle detection *Journal of Marine Science and Engineering* 2023, 11(12): 2363. <https://doi.org/10.3390/jmse11122363>.
 44. Li ZY, Wang Y, Yin ZB, Gao ZB, Wang YJ, Zhen XD. An exploratory numerical study of a diesel/methanol dual-fuel injector: Effects of nozzle number, nozzle diameter and spray spacial angle on a diesel/methanol dual-fuel direct injection engine. *Fuel* 2022; 318(6): 123700. doi:10.1016/j.fuel.2022.123700.
 45. Ke Y, Yao C, Song E, Dong Q, Yang L. Intelligent fault diagnosis method of common rail injector based on composite hierarchical dispersion entropy and improved least squares support vector machine. *Digital Signal Processing* 2021; 114(7): 103054. doi:10.1016/j.dsp.2021.103054.
 46. Dybała J, Mączak J, Radkowski S. *Use of vibroacoustic signal in technical risk analysis*, Institute for Sustainable Technologies, Radom; 2006.
 47. Xu XJ, Yan XP, Sheng CX, Yuan CQ, Xu DL, Yang JB. A belief rule-based expert system for fault diagnosis of marine Diesel engines. *IEEE Transactions on Systems Man and Cybernetics Systems* 2020; 50(2): 656-672. doi:10.1109/TSMC.2017.2759026.
 48. Xu X, Zhao Z, Xu X, Yang J, Chang L, Yan X, Wang G, Machine learning-based wear fault diagnosis for marine diesel engine by fusing multiple data-driven models. *Knowledge-Based Systems* 2020; 190(2): 105324. <https://doi.org/10.1016/j.knosys.2019.105324>.
 49. Bąkowski A., Radziszewski L. *Analysis of selected descriptors of internal combustion engine signals using R software*. Kielce University of Technology, Kielce; 2015.
 50. Zbierski K. *Study of fuel pressure waveforms and atomizer needle movement in injection systems of compression-ignition engines*, Scientific Papers of the Lodz University of Technology, 850, Lodz; 2000.
 51. Monieta J, Kasyk LW. Optimization of the spatial position of the vibration acceleration sensor and the method of determining limit values in the diagnostics of combustion engine injection system. *Sensors* 2026; 26(6): 1981. <https://doi.org/10.3390/s26061981>.
 52. Monieta J, Kasyk L. Application of machine learning to classify the technical condition of marine engine injectors based on experimental vibration displacement parameters. *Energies* 2023; 16: 6898. <https://doi.org/10.3390/en16196898>.
 53. Zieliński TP. *Digital signal processing. From theory to applications*, 2nd ed. Transport and Communication Publishers, Warsaw; 2009.

54. Monieta J, Kasyk L. Optimization of design and technology of injector nozzles in terms of minimizing energy losses on friction in compression ignition engines. *Applied Sciences* 2021; 11(16): 7341. <https://doi.org/10.3390/app11167341>.
55. Jermak C, Majchrowski R, Monieta J. The application of geometric measurement and analysis images for mapping geometric of model marine engine injection equipment. PTNSS Congress, Radom, 16th–17th June 2011, 71-72, cd. 1-8.
56. Balz R., Bernardasci G, von Rotz B, Sedarsky D. Influence of nozzle geometry on spray and combustion characteristics related to large two-stroke engine fuel injection systems. *Fuel* 2021; 294(6): 1-10. doi:10.1016/j.fuel.2021.120455.
57. Kastengren AL, Tilocco FZ, Powell CF, Manin J, Pickett LM, Payri R, Bazyn T. Engine combustion network (ECN): Measurements of nozzle geometry and hydraulic behavior. *Atomization and Sprays* 2012; 22(12): 1011-1052. doi:10.1615/AtomizSpr.2013006309.
58. Marcic M. Measuring method for diesel multihole injection nozzles. *Sensors and Actuators A Physic* 2003; 107(2): 152-158. doi:10.1016/S0924-4247(03)00205-X.
59. Salvador FJ, Gimeno J, de la Morena J., Carreres M. Comparison of Different Techniques for Characterizing the Diesel Injector Internal Dimensions. *Experimental Techniques* 2018; 42(5): 467-472. doi:10.1007/s40799-018-0246-1.
60. Liu B, Gan HB, Chen D, Shu ZP. Research on fault early warning of marine diesel engine based on CNN-BiGRU. *Journal of Marine Science and Engineering* 2022; 11(1): 1-21. doi:10.3390/jmse11010056.
61. StatSoft. *Electronic Statistic Textbook*, StatSoft, 2024; Krakow.
62. Pajak M, Kluczyk M, Muslewski L, Lisjak D, Kolar D. Ship diesel engine fault diagnosis using data science and machine learning. *Electronics* 2023; 12(18): 3860. doi:10.3390/electronics12183860.
63. Montgomery DC, Runger GC. *Applied statistics and probability for engineers*, 1st ed. John Wiley & Sons: Inc., New York; 1994.

Nomenclature

b^*	standardized regression coefficient
b	regression coefficients
B	fuel dose
d	diameter of spray holes
H_i	the i -th amplitudes of the envelope spectra of the vibration displacement signals of the nozzle needle
h	vibration displacement signals of the nozzle needle
h_{max}	maximum needle stroke
i	a number of spray holes
L_a	the average diameter clearance between the mating surfaces
n	camshaft speed
N	fuel rail setting
p	test probability
P_w	injector opening pressure
p_z	injection pump fuel feed pressure
R	regression coefficient
R^2	coefficient of determination
t	student's t test
t_p	fuel temperature at the injection pump inlet
τ	time
α	sealing cone angle of nozzle needle
